Psychological Monographs

General and Applied

No. 368
1953

H. E. Mitzel and W. Rabinowitz

Vol. 67
No. 18

Assessing Social-Emotional Climate in the Classroom by Withall's Technique

By

Harold E. Mitzel and William Rabinowitz

Division of Teacher Education

Municipal Colleges of New York City

Price $1.00

Edited by Herbert S. Conrad

Published by The American Psychological Association, Inc.

Psychological Monographs:
General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology* with the *Psychological Monographs*

Editor

HERBERT S. CONRAD

Department of Health, Education, and Welfare
Office of Education
Washington 25, D.C.

Managing Editor

LORRAINE BOUTHILET

Consulting Editors

DONALD E. BAIER
FRANK A. BEACH
ROBERT G. BERNREUTER
WILLIAM A. BROWNELL
HAROLD E. BURTT
JERRY W. CARTER, JR.
CLYDE H. COOMBS
JOHN G. DARLEY
JOHN F. DASHIELL
EUGENIA HANFMANN
EDNA HEIDBREDER
HAROLD E. JONES
DONALD W. MACKINNON
LORRIN A. RIGGS
CARL R. ROGERS
SAUL ROSENZWEIG
ROSS STAGNER
PERCIVAL M. SYMONDS
JOSEPH TIFFIN
LEDYARD R TUCKER
JOSEPH ZUBIN

MANUSCRIPTS should be sent to the Editor.

Because of lack of space, the *Psychological Monographs* can print only the original or advanced contribution of the author. *Background and bibliographic materials must, in general, be totally excluded* or kept to an irreducible minimum. Statistical tables should be used to present only the most important of the statistical data or evidence.

The first page of the manuscript should contain the title of the paper, the author's name, and his institutional connection (or his city of residence). Acknowledgments should be kept brief, and appear as a *footnote* on the first page. No table of contents need be included. For other directions or suggestions on the preparation of manuscripts, see: CONRAD, H. S. Preparation of manuscripts for publication as monographs. *J. Psychol.*, 1948, 26, 447-459.

CORRESPONDENCE CONCERNING BUSINESS MATTERS (such as author's fees, subscriptions and sales, change of address, etc.) should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D.C.

COPYRIGHT, 1954, BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Vol. 67, No. 18 · Whole No. 368, 1953

Psychological Monographs: General and Applied

Assessing Social-Emotional Climate in the Classroom by Withall's Technique[1]

Harold E. Mitzel and William Rabinowitz
Division of Teacher Education, Municipal Colleges of New York City

I. BACKGROUND OF THE STUDY

IN RECENT years educators have become increasingly concerned with psychological aspects of the classroom environment. It is now generally believed that the goals of education are not maximally achieved unless the social-emotional climate of the classroom is characterized by an atmosphere of warmth, mutual respect, and permissiveness. It is further assumed that the teacher is the most important individual in determining classroom climate, and that her verbal behavior is largely the medium for projecting her influence in the situation.[2]

Based upon these assumptions, Withall (14) has developed a technique for assessing social-emotional climate in the classroom through an analysis of the statements made by the teacher as she conducts classroom activities. It is possible, employing this technique, to place each statement made by a teacher into one of several categories.[3] Starting with an original list of twenty-five categories, Withall found it possible to reduce this number to seven which form a learner-centered to teacher-centered continuum. These seven categories cover all teacher statements in a parsimonious, yet psychologically meaningful fashion. The several categories, and brief examples of each, are as follows:

CATEGORY 1: *Learner-supportive statements* express agreement with ideas, actions, or opinions of the learner, or commend and reassure him. The dominant intent of these statements or questions is to praise, encourage, or bolster the learner. Statements of this kind frequently take the form of a single word response such as "Right," "Yes," or "Good."

CATEGORY 2: *Acceptant or clarifying state-*

[1] It is a pleasure for the authors to acknowledge the considerable assistance and support they have received while carrying out this study. Professor Withall, whose technique was employed, supplied us with sample transcripts and kindly encouraged our efforts. To Mrs. Pauline N. Cagen, Principal of Public School 127 in New York City, we are indebted for unreserved cooperation in providing the observational setting. Mrs. Dora J. Holt, our secretary, provided stenographic assistance in the classrooms and later persevered, with undisturbed good humor, through successive revisions of the manuscript.

To the four teachers who were observed in this study we owe a special debt of gratitude. Being observed necessarily entails certain irritations and frustrations. These they suffered without complaint. The warmth of their reception did much to make this study an enriching personal experience for the authors.

[2] Some recent practice (8, 16) favors the use of the feminine pronoun in referring to teachers, and since only women teachers participated in this study, the feminine gender seems quite appropriate.

[3] A statement is defined in this report as any type of verbalization made by a teacher. Statements require no particular grammatical construction and may consist of questions, declarative sentences, exclamations, and imperatives. In a personal communication from John Withall, dated December 22, 1952, statements are further defined: "The criterion of 'a statement' is the expression of a complete idea or concept . . . any given remark of a teacher may be broken up into two or more statements depending on how many discrete ideas are expressed in said remark."

ments show considerable understanding by the teacher or restate clearly and succinctly in the teacher's words the ideational or emotional content of the learner's statements. The dominant intent of these statements is to help the learner gain insight. Statements of this kind are rather infrequent, but the following is an example: "In other words you think that if we decide to go back again we should plan to spend the entire day."

CATEGORY 3: *Problem-structuring statements* offer facts, ideas, or opinions to the learner about phenomena or procedures in a nonthreatening objective manner. The dominant intent of these statements is to sustain the learner by facilitating his problem-solving activities. A good example of this kind of statement is the following: "What is special about twins? Two brothers are born on the same day one year apart; are they twins?"

CATEGORY 4: *Neutral statements* evidence no supportive intent, being neither teacher-centered, learner-centered, nor problem-centered. The dominant intent of these statements is to facilitate the administrative and procedural detail of class management. Examples of neutral statements are the following: "Children, there will be no movie today since our school projector is broken." "Who is going to march in the St. Patrick's Day parade tomorrow?"

CATEGORY 5: *Directive statements* advise the learner regarding a course of action and offer him no choice or severely limit his choices. The dominant intent of these responses is to have the learner take up the teacher's viewpoint and pursue a particular course of action. An example of a directive statement is the following: "If there are any words which you do not know, put them down in your notebooks and we will go over them tomorrow."

CATEGORY 6: *Reproving, disapproving, or disparaging statements* express complete or partial disapproval of the ideas, behavior, or personal characteristics of the learner. The dominant intent of these statements is to admonish the learner for unacceptable behavior; to instruct the learner in societal values as interpreted by the teacher; or to impress upon the learner that he has failed to reach certain standards of excellence. A good example of mild reproof is the following: "Children, you think you are whispering, but you are not." Stronger reproof is found in this statement: "You knew that yesterday, but now you've forgotten it completely."

CATEGORY 7: *Teacher-supportive statements* are those in which the teacher refers to herself, expressing a defensive attitude, or refers to her present or past interests, activities, or possessions in order to reassure herself and confirm her position. The dominant intent of these teacher responses is to assert, defend, or justify the teacher. An example of this type of statement is the following: "You must say 'excuse me' when you interrupt me; I say 'excuse me' when I interrupt you. Now we will go on without any interruptions."

It is important to note that in using the Withall procedure one categorizes statements not in terms of their *content* but according to the *intent* of the teacher who makes them. Theoretically, the verbal behavior of the teacher, as assessed by the Withall procedure, is uninfluenced by the subject matter taught. In practice, however, it seems likely that such content factors as subject matter do exercise some influence on the frequency of verbal statements in certain categories. The extent of this influence, if any, is not known.

The intensity of a statement is not a variable in the Withall procedure. The technique provides a numerical measure of statements in each category, but the manner in which these statements are made is not assessed. Statements made with great emotion are weighted equally with those made with serene calm or disinterest. This characteristic does not appear to represent a profound problem or limitation, since it is reasonable to expect that intensity of statements and frequency of statements are closely related.

A procedure for categorizing statements necessarily ignores several other facets of communication among human beings. A teacher's posture, gestures, and facial expressions undoubtedly convey meaning to children, and therefore influence the social-emotional climate of the classroom. Even silence may be put to effective use by teachers in the control of classroom situations. Other characteristics of teachers' verbal behavior which conceivably influence classroom climate are the rate and clarity of expression. An essentially dominative social-emotional

climate may be maintained by a teacher who speaks so rapidly and indistinctly that the pupils are insecure in her presence. Though these variables are important factors in the communication between teacher and pupil, they are extremely difficult to assess objectively. Moreover, on logical grounds it would seem that measures of social-emotional climate based upon the categorization of the statements of teachers should be highly correlated with other measures emphasizing different aspects of the verbal pattern.

In his original development of the Climate Index, Withall (14) worked with typescripts made from sound recordings of regular sessions of several high school classes in the University of Chicago Laboratory School. The Index was defined as the ratio of the number of learner-centered statements (sum of Categories 1, 2, and 3) to the number of teacher-centered statements (sum of Categories 5, 6, and 7). Objectivity of the Index was determined upon the basis of agreement among several judges classifying statements in the typescripts. An estimate of validity was obtained by comparing the rank order of three typescripts on the Climate Index with the rank order of Anderson's Integrative-Dominative Ratio (1) using the same three typescripts. Other evidences of validity include a questionnaire evaluation by a class of 25 pupils and the use of several class sessions in which the climate was experimentally varied by the teacher. Reliability was studied by comparing successive blocks of 50 statements from the typescripts of one class session. The extent to which changes occurred in the proportions of statements in each category was noted. Proportions based on the first 200 statements did not significantly change as successive blocks of statements were added. It was therefore concluded that 200 statements offered a reliable sample of a teacher's classroom verbal behavior.

Flanders (5) showed that the Climate Index, when used in an experimental situation with one teacher and one pupil, was related to such physical manifestations of anxiety as galvanic skin response and pulse rate. Using the Index as a criterion of group climate, Perkins (10) categorized the statements of six leaders of in-service teacher training groups. He found that changes in group climate were paralleled, to a marked degree, by changes in group learning.

A further study by Withall (16) had as its purpose the assessment of the classroom climates experienced by a seventh-year class in the Laboratory School as it changed from teacher to teacher. This study emphasized the fact that different teachers create quite different psychological atmospheres for the same group of youngsters. An important change in the procedure for categorizing statements was adopted in this study. Instead of recording class sessions and making categorizations from typescripts, the observer classified teachers' statements in the "live" situation, i.e., as they were made in the classroom.

Thus it can be seen that since its original development, additional research experience has been gained with the Withall technique. Applications have ranged from junior high school classrooms to adult groups enrolled for in-service teacher training, from a "class" involving a single pupil to classes composed of large numbers of pupils. Although used exclusively in educational settings thus far, the method undoubtedly is applicable in the study of psychotherapy, counseling, group discussion, industrial interviewing, and other human relationships in which interaction takes place mainly on a verbal level.

II. PROBLEM

The review of the literature cited above indicated the desirability of answering certain basic questions concerning the usefulness of a verbal-behavior method in studies of teacher effectiveness. Withall's technique seemed suitable as a predictor or criterion of teacher effectiveness (9, 11), but its immediate use in this capacity appeared to depend upon its reliability. Accordingly, a study was planned which might answer three fundamental questions, usually subsumed under the general heading of reliability. These questions are as follows:

a. Can self-trained observers, working independently and observing the same

behavior, agree in their classification of teachers' statements?

b. Is the verbal behavior of teachers stable from one classroom visit to another?

c. In a single school, is the categorization of verbal behavior sufficiently sensitive to reveal differences in the classroom climates provided by different teachers?

A satisfactory answer to each of these questions seems preliminary to the design of studies of teacher effectiveness employing the technique.

III. PROCEDURE

Exploratory Efforts

The initial efforts of the authors were directed toward a clarification of the meaning of the various categories. It was essential that the authors have a common understanding concerning the types of teacher statements which properly fit into each category. Professor Withall, in a private communication, encouraged the authors to employ the method and provided them with several sample typescripts as practice material. The principal of a small Manhattan elementary school, with the consent and approval of her 13 faculty members, agreed to furnish the observational situation and to make minor adjustments in the school schedule for the convenience of the observers. At aperiodic sessions during a six-week interval, classroom visitation and practice with the classification scheme were undertaken. It soon became apparent that sound reproductions of class sessions using commercial tape recorders were not feasible because of the high noise concentration in the classrooms. Stenographic notes of teachers' verbalizations were found to be more satisfactory than tape recordings, but even with this method the authors estimated that no more than 75 per cent of a teacher's statements were obtained during any given observation period. After about 20 clock hours of experience in classifying classroom verbal behavior of teachers from typescripts (see Appendix), the writers attempted to dispense with reproduction procedures and to categorize statements in the "live" situation. As a teacher conducted a class, the observers (the writers) independently categorized her statements. Practice for a few sessions with this method seemed to yield close agreement between the observers as to the total number of statements made by the teacher and the proportionate numbers of statements in the several categories.

Observational Setting

The school in which the observations were made is somewhat atypical for the country as a whole. It is a small, six-year elementary school, housed in a 100-year-old building, and located in an economically underprivileged area of New York City. Approximately 30 children are enrolled in each class. Many of the pupils are foreign born; about one-half are natives of Puerto Rico. Most of the children are bilingual. Measured IQ's are almost all below 100, but this is certainly attributable, in part, to difficulties in the use of the English language.

From an impressionistic point of view the atmosphere of most classrooms in this school was free and informal. The authority of the teacher, though expressed in subtle rather than obvious ways, was seldom challenged. Disciplinary problems were infrequent but not unknown. The teachers, if one may characterize a group of teachers in terms such as these,

are neither "traditional" nor "modern" but at some indeterminate point between these two extremes. Such practices as severe regimentation and repetitive drill were never seen; neither was there evidence of the studied encouragement of pupil self-direction.

It is impossible to estimate the extent to which these teachers are representative of teachers in general. Like most New York City teachers, they are well educated, with a strong, professional interest in their work. Their original attitude toward the study could best be characterized as "cautious toleration." During the course of the observations, however, this attitude underwent a marked change. Rapport was established as the observers became more familiar and less threatening figures. The teachers gave definite evidence of enthusiastic interest and unreserved cooperation as the study progressed.

Design

Three important aspects of the categorization of teachers' verbal behavior were the principal objects of study: (*a*) differences between observers, (*b*) differences among observation periods or visits, and (*c*) differences among teachers. A $2 \times 4 \times 4$ multiple-classification design involving two observers, four visits, and four teachers was employed. This was considered a first series of observations. It was followed by a second series of observations, of identical design, in which the same observers visited the same teachers on four more occasions. Dividing the eight observation days into two separately analyzed series of four visits each enabled the observers, after the first series, to re-evaluate their criteria for classifying statements into the several categories. It was thought that with this method the categorizations by observers could perhaps be brought into closer alignment for the second series of four visits.

The data were to be treated by the analysis of variance method, and each category of statements was to be analyzed separately. During the exploratory work, however, it became apparent that statements in Category 2 (acceptant or clarifying) and Category 7 (teacher-supportive) were very rare. It was therefore decided that analysis should be restricted to each of the five "most tallied" categories and a summation category (Climate Index).

Two fourth-year and two fifth-year teachers in the cooperating school were chosen for study. None of the classrooms of these teachers had been visited during the practice and exploratory sessions, and the observers had no knowledge of the methods and techniques employed by these teachers prior to the first observation. Observations were held on Monday mornings for eight consecutive weeks, exclusive of school holidays. Each teacher was visited some time during the morning. The order of visitation of classrooms can best be described as haphazard. Since time of day may conceivably be related to classroom climate, it was desirable to avoid a systematic order. At the same time it was impossible, because of scheduling difficulties, to utilize a randomizing technique. On two occasions, as a result of absence, it was not possible to visit a particular teacher on Monday morning, and a later day in the same week was chosen as a substitute.

Previous experience with Withall's procedure suggested that approximately 100 statements would provide a usable sample of a teacher's verbal behavior for a single observation period. Withall (15) concluded that a sample of 50 statements was sufficient for bringing out gross dif-

ferences in classroom climate and that 200 statements were sufficient for estimating a teacher's position along the climate continuum. However, since it seemed desirable for the two observers to impose upon a given teacher no longer than 30 minutes per day, a goal of approximately 100 teacher statements per visit was selected. In practice the number of statements tallied by the two observers on the 32 visits of both series varied from 74 to 133 per visit, with a median of 100 statements. Data relative to this point are presented in Tables 1 and 2.

Factors not studied. Ultimately, extensive evidence on the validity of Withall's procedure for categorizing teacher statements will have to be obtained. The pupil-growth correlates of a teacher's verbal behavior must be sought. The discovery of such correlates, however, is not within the scope of this study, nor does this study deal with an assessment, inference, or assumption concerning the effectiveness of the teachers who served as subjects.

It should be noted that several factors which may ultimately prove to influence observed classroom climate have been uncontrolled in this study. Factors of this type would include the day of the week, the time of day, and the background and experiences of the pupils. It was the feeling of the writers that, though these factors are worth studying, they contribute a small proportion to the variation in observed classroom climate.

Assumptions. Several assumptions are required for the multiple-classification design and analysis of variance technique to be strictly applicable. Cochran (3) has set forth these assumptions as independence, homogeneity of variance, and normality of distribution of the experimental errors, as well as additivity of treatment effects. Concerning this study, which was to involve the repeated observation of the same teachers by the same observers, it seemed that the independence of errors was most in jeopardy. To minimize this difficulty, special care was taken by the observers during each series of visits to avoid any exchange of impressions about the teachers. Also, by placing observation records in a file and by carefully refraining from totaling the tally marks in each category following an observation period, each observer was able in a measure to prevent contamination of his subsequent categorizations of statements. Since inspection of the data provided no evidence for heterogeneity, the assumption of homogeneous error variance was considered satisfied without rigorous test.

Nature of the variate. Following the precedent set by Withall (15) in his original development of the Social-Emotional Climate Index, it was decided to make the entry for each cell of the design the percentage of the total number of a teacher's statements falling in a given category. This procedure has the effect of equating all observation periods to a base of 100 teacher statements, making it possible to compare varying numbers of statements recorded by different observers during different visits. Distributions of percentages, however, are characteristically rectangular in form and transformation is required to normalize these values. Bartlett (2) points out the usefulness of the angular transformation for percentages and suggests the employment of Table XII in Fisher and Yates's *Statistical Tables* (4). This method was followed for these data. For example, if Teacher A has a total of 99 statements attributed to her by Observer 1 during a given visit, and 25 of these statements have been classified by him as "learner-supportive" (Withall's Category 1), then 25.3 per cent of her verbal behavior is placed in the first category. Using the angular transformation table, 25.3 per cent is transformed into an angle of 30 degrees, rounded to the nearest whole degree. In this way seven transformed values, one for each statement category,

were obtained for each of four teachers, by each of two observers, on each of four visits, in two separate series of observations.

As anticipated, examination of the tables of transformed values indicated that it would be worth while to analyze only five of the seven statement categories and a Climate Index based on the sum of the first three categories. So few teacher statements were classified by the observers as belonging in Category 2 and Category 7 that an analysis of these was not undertaken.

The Index used in this study was the sum of statements in Categories 1, 2, and 3 (learner-centered statements) divided by the total number of statements. This Index was computed separately for each visit to each teacher by each observer. There were therefore 32 indices computed for each series of observations. It should be noted that the Index used in this study, since it is based upon proportions of the total number of statements, will vary between zero and unity (before transformation). Withall's ratio of learner-centered statements to teacher-centered statements, ignoring neutral statements (Category 4), will vary between zero and some indeterminate value, depending upon the number of statements categorized by the observer.

IV. ANALYSIS OF RESULTS

Comparisons of Numbers of Statements Recorded by Observers

One of the preliminary questions for study in this project was, "Can two persons, observing simultaneously and independently in the classroom, agree upon the total number of statements made by a teacher?" Table 1 provides information relevant to this question. During the first series of observations there was some tendency for Observer 2 to record more statements than Observer 1–1,741 compared with 1,693. Differences between observers on number of tallied statements for the 16 pairs of observations (four visits to four teachers) resulted in a t value of 2.216, which is significant at the 5 per cent level ($t_{.05}=2.131$; 15 df). Thus, during the first series of observations one observer tended to record a significantly larger number of statements than the other. This difference, although undesirable, does not necessarily lead to disagreement between observers when statements in individual categories are analyzed. Statements were analyzed for each observer in the separate categories by calculating a *ratio* based upon the total number of statements tallied during the observation period. Two ratios may be identical though the total number of statements on which they are based differ markedly.

Between the first and second series of observations the observers redefined their criteria of a statement in order to reach greater agreement. The success of their effort is reflected in Table 2. The mean difference for the 16 pairs of observations of this second series resulted in a t value of 0.260, which is nonsignificant. Thus, during the second series of observations, after the observers re-evaluated their criteria, the systematic significant difference in the number of recorded statements was eliminated.

Comparison of Range and Median Climate Indices

Table 3 presents the Social-Emotional Climate Indices of the four teachers who

TABLE 1

TOTAL NUMBER OF STATEMENTS AND LENGTH OF RECORDED OBSERVATION PERIODS

First Series

Visit	Obs.	Teacher A		B		C		D		Totals	
		N	Time	*N*	Time	*N*	Time	*N*	Time	*N*	Time
I	1	105	20′	117	16′	105	20′	137	20′	464	76′
	2	109		126		116		138		489	
II	1	95	24′	124	13′	109	23′	89	12′	417	72′
	2	93		123		111		104		431	
III	1	78	21′	106	24′	110	30′	102	17′	396	92′
	2	76		101		111		107		395	
IV	1	93	18′	123	22′	84	27′	116	22′	416	89′
	2	96		120		87		123		426	
Totals	1	371	83′	470	75′	408	100′	444	71′	1693	329′
	2	374		470		425		472		1741	
Medians	1	94	20.5′	120	19.0′	107	25.0′	109	18.5′		
	2	95		122		111		115			

TABLE 2

TOTAL NUMBER OF STATEMENTS AND LENGTH OF RECORDED OBSERVATION PERIODS

Second Series

Visit	Obs.	Teacher A		B		C		D		Totals	
		N	Time	*N*	Time	*N*	Time	*N*	Time	*N*	Time
I	1	99	16′	95	11′	74	25′	88	15′	356	67′
	2	100		92		79		93		364	
II	1	93	25′	87	19′	92	30′	87	28′	359	102′
	2	85		85		95		85		350	
III	1	120	23′	133	23′	91	35′	100	18′	444	99′
	2	108		132		89		106		435	
IV	1	101	20′	78	20′	94	32′	92	22′	365	94′
	2	102		86		104		89		381	
Totals	1	413	84′	393	73′	351	122′	367	83′	1524	362′
	2	395		395		367		373		1530	
Medians	1	100	21.5′	91	19.5′	91	31.0′	90	20.0′		
	2	101		89		92		91			

participated in the study for all observation sessions. A separate analysis of the observations made in the two series is not provided at this point, since that is considered in other sections of this report. The tabular values are based upon the following formula:

$$\frac{\text{Sum of Statements in Categories 1, 2, and 3}}{\text{All Recorded Statements}} \times 100$$

The three categories of the numerator (learner-supportive, acceptant or clarifying, and problem-structuring) may be considered learner-centered statements. Their ratio to the total number of statements made by the teacher is the classroom Climate Index employed in this study. Indices for the two observers were combined and averaged, since differences between the observers were generally not significant, as subsequent analysis will demonstrate.

Examination of the median values near the bottom of Table 3 shows that Teacher A, with almost 60 per cent of her statements in the learner-centered categories, has the highest Climate Index. Teacher C, on the other hand, has the lowest Index, 33.9 per cent. Teachers B and D have nearly equal Climate Indices, based on slightly less than one-half of their statements in the learner-centered categories. It was the subjective impression of the observers, prior to the analysis of the data, that the four teachers tended to provide similar classroom climates. A larger sampling of the teaching profession would undoubtedly yield a wider range of individual differences in average Climate Indices.

TABLE 3

SOCIAL-EMOTIONAL CLIMATE INDICES (PERCENTAGES) FOR FOUR TEACHERS ON EIGHT VISITS

Visit	Teacher			
	A	B	C	D
1	58.4	26.5	44.5	47.6
2	61.1	53.9	35.0	33.9
3	42.9	35.3	37.1	56.5
4	82.0	47.0	25.8	40.6
5	54.8	65.3	43.1	55.9
6	74.6	41.3	32.7	50.0
7	48.2	40.4	24.5	27.7
8	78.3	66.5	31.3	41.4
Median	59.8	44.2	33.9	44.5
Range	39.1	40.0	20.0	28.2

Table 3 not only shows differences among teachers in typical classroom climate but differences in the climate provided by the same teachers on different occasions. A picture of the variability in the climates characteristic of the four teachers is obtained by considering the range of percentage points covered by the Indices of each teacher for the eight observations. Teachers A and B are of approximately equal variability and are relatively more variable than Teachers C and D. These data suggest that not only do teachers differ among themselves on their Climate Indices, but they also differ in the consistency of classroom climate which they provide from occasion to occasion. These impressions are gained from an inspection of Table 3 and will be analyzed more thoroughly in the material to follow.

Results and Interpretation of Analyses of Variance

Table 4 shows the structure of the study for the first series of four observations, and Table 5 is a similar presentation for the second series of four visits. In all cases the second-order interaction ($T \times O \times V$) was used as the error term for testing lower-order effects. Results are shown separately for five categories of teacher statements and the Climate Index based on the sum of the statements in the first three categories.

The interpretations suggested in this

TABLE 4

SUMMARY OF ANALYSES OF VARIANCE FOR FIVE CATEGORIES OF STATEMENTS AND AN INDEX OF TEACHERS' VERBAL BEHAVIOR

First Series

	Source of Variation							
	Main Effects			Interactions				Total Mean Square
				1st Order			2nd Order	
	Among Teachers	Among Visits	Between Observers	T×O	T×V	O×V	T×O×V (Error)	
Degrees of freedom	3	3	1	3	9	3	9	31
Category 1 Learner-supportive								
Mean Square	27.542	29.708	0.125	2.875	33.792	7.708	1.236	16.738
F	22.28**	24.04**	<1.00	2.33	27.34**	6.24*	—	—
Category 3 Problem-structuring								
Mean Square	301.135	3.125	6.125	2.458	91.347	0.458	4.458	57.738
F	67.55**	<1.00	1.37	<1.00	20.49**	<1.00	—	—
Category 4 Neutral								
Mean Square	7.865	82.865	22.781	17.531	29.670	16.365	3.448	22.410
F	2.28	24.03**	6.61*	5.08*	8.60**	4.75*	—	—
Category 5 Directive								
Mean Square	339.583	44.250	4.500	1.917	99.000	6.417	2.833	67.661
F	119.87**	15.62**	1.59	<1.00	34.95**	2.27	—	—
Category 6 Reproving								
Mean Square	38.458	48.875	6.125	1.125	29.431	0.708	1.819	17.899
F	21.14**	26.87**	3.37	<1.00	16.18**	<1.00	—	—
Climate Index								
Mean Square	348.792	22.042	3.125	4.792	127.819	5.375	4.597	75.415
F	75.87**	4.79*	<1.00	1.04	27.80**	1.17	—	—
Tabled *F* Values								
$F_{.01}$	6.99	6.99	10.56	6.99	5.35	6.99		
$F_{.05}$	3.86	3.86	5.12	3.86	3.18	3.86		

* Exceeds the 5 per cent level of significance.
** Exceeds the 1 per cent level of significance.

presentation and discussion of results are characteristic of the four teachers studied and can be generalized only to whatever hypothetical population they represent. As in much educational research, it was not possible to select the particular elementary school at random from a population of New York City schools. Nor was it possible, for administrative reasons, to select randomly four teachers from among the particular school's 13 faculty members. There is, therefore, no reason to believe that the four teachers in this investigation are representative of teachers generally. In the future, if it is desirable to estimate parameters for the Climate Index or to construct norms, considerable attention will have to be given to problems of sampling.

As in the case of the teachers, the observers also constitute a sample from a hypothetical population of observers. Although the categories are objectively defined, it is possible that observers with radically different orientations might categorize the same teachers' statements in somewhat different ways.

Main effects. Of the three main effects which are of concern in this study, observers, visits, and teachers, the source of variation due to the observers is of initial importance. The third column in Tables 4 and 5 presents the between-observers mean square and appropriate *F* test for

TABLE 5

SUMMARY OF ANALYSES OF VARIANCE FOR FIVE CATEGORIES OF STATEMENTS AND AN INDEX OF TEACHERS' VERBAL BEHAVIOR

Second Series

	Source of Variation							
	Main Effects			Interactions				Total Mean Square
				1st Order			2nd Order	
	Among Teachers	Among Visits	Between Observers	T×O	T×V	O×V	T×O×V (error)	
Degrees of freedom	3	3	1	3	9	3	9	31
Category 1 Learner-supportive								
Mean Square	69.615	181.031	0.281	0.281	17.920	1.365	0.698	29.830
F	99.73**	259.36**	<1.00	<1.00	25.67**	1.96	—	—
Category 3 Problem-structuring								
Mean Square	679.781	46.615	2.531	2.281	116.976	2.615	2.698	105.596
F	251.96**	17.28**	<1.00	<1.00	43.36**	<1.00	—	—
Category 4 Neutral								
Mean Square	52.375	28.208	24.500	8.083	36.125	12.750	7.000	23.125
F	7.48**	4.03*	3.50	1.15	5.16*	1.82	—	—
Category 5 Directive								
Mean Square	664.365	186.781	1.531	1.031	115.781	3.615	2.226	117.128
F	298.46**	83.91**	<1.00	<1.00	52.01**	1.62	—	—
Category 6 Reproving								
Mean Square	76.115	28.698	13.781	0.365	44.003	1.115	3.809	24.612
F	19.98**	7.53**	3.62	<1.00	11.55**	<1.00	—	—
Climate Index								
Mean Square	498.531	236.948	3.781	1.865	70.920	1.281	2.920	93.039
F	170.73**	81.15**	1.29	<1.00	24.29**	<1.00	—	—
Tabled *F* Values								
$F_{.01}$	6.99	6.99	10.56	6.99	5.35	6.99		
$F_{.05}$	3.86	3.86	5.12	3.86	3.18	3.86		

* Exceeds the 5 per cent level of significance.
** Exceeds the 1 per cent level of significance.

each category of teachers' statements in both series of observations. In the first series of visits differences between the two observers were attributable to chance variation in all categories except neutral statements, Category 4. Here the issue remained in doubt since the *F* test yielded significance between the 5 and 1 per cent levels. An examination of the means showed that Observer 1, on the average, placed 15 per cent of all teachers' statements in this category, while Observer 2 similarly categorized 13 per cent from the same set of teachers and visits. In the analysis of the second series of visits, shown in Table 5, the main effect due to observers is nonsignificant in all categories, including the category for neutral statements.

The remaining main effects, among-visits and among-teachers, are interesting, but were not the chief concern of the investigation. Most of the analyses reported in Tables 4 and 5 show highly significant results for the among-teachers and the among-visits sources of variation. These results indicate that random sampling variation fails to account for the observed variability in these sources. It should be emphasized, however, that we cannot say anything further regarding a population of visits or a population of teachers. The existence of a significant interaction between teachers and visits

(T × V), interpreted as being characteristic of the population, prevents us from making inferences regarding the "real" effects of either teachers or visits considered separately. At the same time we are prevented from using the teacher-by-visit interaction as an error term. Since it is significant, it does not provide an appropriate estimate of error; consequently, no meaningful null hypothesis is available to test. In nonstatistical language, the variation among teachers on different visits is of such magnitude that the variability among either teachers or visits is not made apparent for the majority of statement categories. This point will be explored further in the discussion of the interactions of the analysis of variance tables.

Interactions. Considering the first-order interactions listed in Tables 4 and 5, the general picture is one of nonsignificant effects for those interactions involving observers. The teacher-by-visit interaction with its nine degrees of freedom is in a sense the primary focus of the study of interactions. As indicated in the discussion of the main effects, this term (column 5 of Tables 4 and 5) was found to be significant for all of the categories on which the analysis was made. It shows that teachers are quite variable from time to time with respect to the proportions of their verbal behavior occurring in the categories of Withall's technique.

The extent to which differences among teachers, in general the largest single source of variation in the data, are masked by the variation due to the interaction of teacher-by-visit is an important problem. A rule-of-thumb procedure for examining this point consists of the formation of a ratio, the among-teachers mean square over the teacher-by-visit interaction mean square. Lindquist (7, p. 214) suggests that this ratio of variances is seldom amenable to the F test for significance because of the inability to postulate a meaningful hypothesis. The ratio does, however, when it is several times unity, provide an accurate notion of the stability of the respective rank orders of the elements being studied, in this case teachers. An example from Table 4 will make this point clearer. The ratio of the among-teachers mean square and the teacher-by-visit mean square for the Climate Index in the first series of observations is:

$$348.792/127.819 = 2.73$$

This ratio is not very large, and we would therefore not expect great stability in the rank orders of the four teachers over the four visits. If the data for the two observers are combined, the set of rank orders is given in Table 6. Using a nonparametric method recommended by Kendall (6), a coefficient of concordance computed from the ranked data in Table 6 was found to be .475. This value is nonsignificant, and it would therefore appear that the distribution of ranks in the Climate Index for the first series of observations might easily have arisen by chance.

TABLE 6

TEACHERS' RANK ORDERS ON CLIMATE INDEX FOR FOUR CLASS VISITS

First Series

Teacher	Visit			
	I	II	III	IV
A	1	1	2	1
B	4	2	4	2
C	3	3	3	4
D	2	4	1	3

An identical analysis for the second series of observations yielded somewhat different results. The ratio of the among-teachers mean square and the teacher-by-visit mean square for the Climate Index

for this second series of observations is:

$$498.531/70.920 = 7.03$$

This ratio is much larger than unity, and, in addition, much larger than the comparable ratio for the first series of observations. If the data are combined for the two observers on the second series of visits, the set of rank orders of the four teachers on each of four visits is given in

TABLE 7

TEACHERS' RANK ORDERS ON CLIMATE INDEX FOR FOUR CLASS VISITS

Second Series

Teacher	Visit			
	I	II	III	IV
A	3	1	1	1
B	1	3	2	2
C	4	4	4	4
D	2	2	3	3

Table 7. It can be seen that the agreement among this set of ranks is quite close and that the average order for the teachers would be A, B, D, and C. The coefficient of concordance was found to be .700. The likelihood that the arrangement of ranks presented in Table 7 could occur by chance if there were no fundamental agreement among them is less than two in one hundred. Thus it appears possible, in spite of fluctuations from occasion to occasion, to order the four teachers participating in this study with respect to their standing on the Climate Index. The stability of ranks necessary for ordering the teachers did not become apparent, however, until the second series of observations. These findings, if they can be duplicated in other school situations among other teachers, suggest that the Climate Index represents a stable variable in spite of fluctuations from time to time for the same teacher. As in this study, stability of ranks may not become apparent until a large number of observations have been made. In school situations where teachers differ markedly, fewer observations will be required in order to demonstrate stability among ranks.

Comparison of the First and Second Series of Observations

In general, although the two series of observations were comparable in virtually every respect, certain unanticipated differences between them appeared. The results reported in Table 5 for the second series of visits tended to confirm the outcomes of the first series reported in Table 4, i.e., significance tended to be characteristic of the same main effects and interactions. Examination of the two tables does, however, make it possible to suggest additional outcomes. There was a general tendency for the total variances of the six categories (last column) to be larger for the second series of observations. Along with this general increase in variability there was a tendency for the error terms to be smaller in the second series than in the first, with the exception of Categories 4 and 6. These effects corroborated two subjective notions entertained by the observers:

a. Practice effects for making categorizations of teachers' statements were believed to have occurred. The observers felt that, with continued practice and the discussions between the first and second series of observations, they became more precise in their ability to recognize and categorize statements. Increased skill in the recognition and categorization of teachers' statements seems to offer a reasonable explanation for the slight decreases in most error terms from the first to the second series.

b. Teachers' verbal behavior was markedly influenced by the presence of the

observers in the classroom. Initially, the teachers seemed to show an undue awareness of the observers. The observers got the impression, during the first few visits, that they were witnessing "tailor-made" lessons. Both the content and the manner of presentation seemed to have been especially prepared for the observers' benefit. As repeated visits were made to the same classrooms, it was felt that this rather obvious planning became much less important. During the later visits the teachers seemed to provide a more typical, less guarded classroom performance. On logical grounds, it appears likely that the net effect of "preparation for being observed" is to reduce the differences among teachers in all statement categories and to have a general depressing effect upon the total variance for each category. Though it cannot be proved conclusively, this is a plausible explanation for the increase in the among-teachers mean square and in the total variances from the first series of visits to the second.

V. SUMMARY AND CONCLUSIONS

This study was designed to investigate three aspects of the reliability of teachers' verbal behavior. The specific questions on which the project was focused are as follows:

a. Can self-trained observers, working independently and observing the same behavior, agree in their classification of teachers' statements?

b. Is the verbal behavior of teachers stable from one classroom visit to another?

c. In a single school, is the categorization of verbal behavior sufficiently sensitive to reveal differences in the classroom climates provided by different teachers?

The method developed by Withall (13, 14, 15) for assessing classroom climate by classifying verbal behavior along a learner-centered to teacher-centered continuum was employed. Using Withall's seven categories for tallying teachers' statements, a multiple-classification design was developed in which two observers (the writers) visited the classrooms of four teachers in an elementary school on four different occasions approximately one week apart. This design constituted the first series of observations and was followed by a second series of identical design involving the same teachers and observers.

In analyzing the data the first problem considered was the extent to which the two observers agreed on the total number of statements uttered by the teachers during each observation. In the first series of 16 observations a statistically significant difference (5 per cent level) between the observers was found. One observer tended to record a slightly larger number of statements than the other. Between the first and second series of observations the two observers attempted to reach greater agreement on the defining characteristics of a statement. This effort was apparently successful, since no significant difference in the number of recorded statements was found on the second series of observations.

A Climate Index based upon the proportion of learner-centered statements to all recorded statements was calculated for the four teachers on the eight observation sessions. A descriptive presentation of these indices indicated that teachers vary in the typical classroom climate they provide. In addition, teachers appeared to differ in the consistency of classroom climate which they provide from one occasion to another.

The data were further treated by

means of the analysis of variance, five statement categories and the Climate Index being analyzed separately. The most important outcomes of this analysis were:

a. In the first series of visits differences between the two observers were attributable to chance variations in all categories except neutral statements, Category 4. In the second series of visits, however, the main effect due to observers was nonsignificant in all categories, including the category for neutral statements.

b. The factors accounted for by differences among the observed teachers and among visits were attributable to nonrandom sources. Since the teacher-by-visit interaction term was found to be significant, no further interpretation regarding these main effects could be made.

c. The significant effect for teacher-by-visit is an answer to the questions posed in this study concerning differences among teachers and stability of teachers' verbal behavior from time to time. It was concluded that teachers showed marked fluctuations in their verbal behavior; "real" differences among teachers were apparent in spite of these fluctuations. It was not until the second series of observations, however, that the teacher differences were clearly manifest.

In the comparison of the results of the first and second series of observations, two impressions stood out. The observers were able to increase their skill in categorizing teachers' statements during the course of the experiment, and teachers' verbal behavior was markedly influenced by the presence of the observers in the classroom, particularly during the early visits.

Concerning the generalizability of these findings it should be noted that the observed teachers are not representative of teachers in general. They were elementary teachers in a particular school situation in the largest city in the nation. Had it been the writers' purpose to estimate the typical classroom climates provided by teachers in the country at large, then the data would be seriously inadequate. Since this was not the goal, the fact of unrepresentativeness is less damaging. A second qualification involves the fact that samples of approximately 100 teacher statements were tallied by the observers during any given visit. There is no way of inferring from the data in this report whether statement samples smaller or larger than 100 would yield comparable results.

In the opinion of the writers several extensions of this research could be undertaken with profit:

a. The validity of the Climate Index needs to be pursued in terms of several criteria of teaching effectiveness, particularly pupil growth. In the absence of wholly adequate criteria it is important to determine some of the correlates of teachers' verbal behavior. Some of these factors could be subsumed under teacher personality and experiential background. Others might involve responses of classes and pupils and characteristics of the school situation.

b. Normative studies for determining the range of verbal behavior and classroom climate among types of teachers and among types of schools would be desirable. In this connection special attention must be paid to sampling procedures.

c. The consistency of the climate provided by a teacher may be an aspect of classroom behavior worthy of intensive study. It may be found that whether a teacher's average Climate Index is 25 or 75 is not as important as whether she tends to create a consistent atmosphere in which youngsters know what to expect. It can be hypothesized that consistency of classroom climate reduces anxieties on the part of pupils, and that such anxiety reduction facilitates social and intellectual growth.

d. Another extension, which should perhaps precede the others, involves the determination of the size of the most appropriate statement sample for picturing a teacher's verbal behavior. The arbitrary sample size of 100 statements employed in this study may account for some of the inconsistency in teachers' verbal behaviors from occasion to occasion. A study needs to be made in which varying numbers of statements are recorded and compared.

APPENDIX

Although the procedures involved in assessing classroom climate by categorizing the statements of the teacher have been described in the body of this report, the reader who wishes to utilize this method for research purposes will probably feel the need for a more thorough discussion. The material to be presented in this appendix is designed to satisfy this need, in part. For the most complete treatment, the reader is urged to consult articles by Withall (15, 16) and, in particular, his thesis (14).

A Suggested Procedure for Using Withall's Technique

There is no single procedure for developing skill in the categorization of verbal behavior which will prove best under all circumstances. The problem under investigation and the resources available to the investigators will, for the most part, suggest efficient procedures. It seems possible, however, to list certain principles which, under virtually all circumstances, research workers may adopt with profit. Much of the material to follow has been adapted from Withall's thesis, as modified in the light of the experience of the present authors.

Plan to work in teams, not alone. For many reasons it is wise in attempting to employ Withall's technique to work with others rather than alone. Only when categorizations have been made by at least two persons, working independently, is it possible to obtain a measure of agreement between judges. Moreover, in the early stages of practice with the technique, careful consideration of those statements on which disagreement is found will prove especially helpful in clarifying the criteria for the various categories. Needless to say, after considerable practice with the technique and satisfactory evidence of agreement with other judges, an individual judge may undertake categorizations alone.

Review the literature on the technique. A considered review of the literature will not only reveal the various problems to which Withall's technique has been applied, but will also indicate ways in which the data obtained by this technique may be treated statistically. An especially careful review should be made of the criteria for the seven statement categories. Each judge should attempt to develop a thorough understanding of what constitutes a statement and the defining properties of the various categories. Discussion among the judges at this stage will help prevent misconceptions and will tend to produce more uniform understandings of the category criteria.

Employ practice transcripts. Transcripts of class sessions should be obtained and used by judges as practice material. Withall's thesis contains a short practice transcript, and the two transcripts which follow may be used for practice purposes by masking the categorizations given each statement. All categorizations should be done independently so as to avoid contamination. Comparisons of the categorizations of different judges will tend to reveal significant areas of disagreement which will require restudy of category criteria. In the initial stages of skill development it will frequently be necessary to refer to the criteria of the seven categories as one actually categorizes a transcript. At later stages, however, this will prove less necessary and should be avoided.

Observe in schools. As soon as judges have developed some skill in categorizations, classroom observations in which verbatim records are made should be undertaken. As before, categorization should be based upon the transcribed record, but the categorizing process will present fewer difficulties since the judges were present during the actual class session. Discussions between observers will introduce additional clarifications.

Categorization may be attempted in the "live" situation. Under certain circumstances it may prove desirable to avoid any procedure for taking a verbatim record of what the classroom teacher says, and instead categorize her statements as she makes them, i.e., in the "live" situation. This should not be attempted unless the observers are thoroughly acquainted with the Withall technique, since it requires the ability to recognize and categorize statements with great rapidity. The advantage of this procedure lies partly in its simplicity, for it requires neither expensive, temperamental, sound-recording equipment nor a highly skilled stenographer. At the same time, by categorizing statements as they are made, the observer gains additional cues from the teacher's posture, gestures, facial expressions, and the like.

The presence of sound equipment or a recording stenographer in the classroom is often disturbing to the teacher and class. Under such circumstances a representative sample of a teacher's classroom behavior is difficult to obtain. In certain situations teachers or administrators will not permit verbatim records to be made. In other situations the presence of a stenographer or recording apparatus in the classroom will indicate to the teacher that her *verbal* behavior is of major concern to the investigators.

For these reasons the present writers recommend that Withall's technique be employed in the "live" situation, if the observers have sufficient skill to do so. The principal disadvantage of this method is that no permanent, accurate record of what was actually said by the teacher is available. This disadvantage will usually be

minor, however, and may be partially overcome by having the observers write a brief description of the class session after it has transpired.

Transcriptions of Actual Class Sessions

Presented below are two transcriptions of portions of actual class sessions. The first, Transcript A, is taken with Professor Withall's permission from his thesis (14). Transcript B was obtained by the authors of this report in the early stages of their practice observations.

TRANSCRIPT A

(NOTE: Three diagonal lines at the end of a statement denote an interruption of the person speaking.)

Teacher C is giving individual help in a Grade VII mathematics class.

		Category
1.	4 C:	So many people are absent today that I question whether we should er work ahead on er anything else of ///
	Pupil:	Let's finish our graphs.
2.	5 C:	I think the thing that you'd better do is, first, if you haven't finished these tables on heights and weights you'd better do those, both height and weight and the graph if you haven't finished that. If you have finished those, if you've finished the graphs, then, both the tables and the graphs, then ///
	Pupil:	Miss C.?
3.	4 C:	Yes?
	Pupil:	Do we have to hand the weight in? I haven't finished height.
4.	3 C:	No! I'm giving you time now to finish both tables, if you've finished both of those. . . .
	Pupil:	Miss C!
5.	4 C:	Yes?
	Pupil:	(indistinguishable)
6.	(*a*) 3 C:	(*a*) Those are the graphs? I see.
	(*b*) 7	You were told to keep those until we can talk about them.
	(*c*) 6	(*b*) I intended to start the discussion of them today but so many are away that I think we hadn't better, uh, everybody. (*c*) Judy!! I've just said what? Huh?
	Judy:	I didn't hear . . . (indistinguishable)
7.	3 C:	I said we could finish them right now. If everybody has finished both the tables and the graphs.
	Pupil:	Miss C, do we have to hand in both height and weight?
8.	3 C:	Eventually.
	Pupil:	Not today, we don't have to hand them in today.
9.	6 C:	No we're finishing them—you were absent, so you'll have to wait until I can explain to you what you are supposed to be doing.
	Pupil:	Shall I explain it to him?
10.	(*a*) 1 C:	(*a*) You can, if you can do it
	(*b*) 5	quietly. (*b*) Just a moment I think that if, er, if you've finished all this we've, we've done most of the work up to page 163 in the book, those who have finished all of this can start to study page 163 and start on that list of problems there. See?
	Pupil:	(indistinguishable)
11.	(*a*) 3 C:	(*a*) No I want them finished and
	(*b*) 1	then you're going to tell us about them orally. Just as soon as you've finished, maybe Friday, maybe not 'til Monday. See? But you'll get them done so that you'll have a story to tell. You see? (Pupil: Yah.) (*b*) All right. (No talking to be heard from anyone for fifteen to twenty seconds.)
12.	5 C:	All right let's get these done now and if you've finished all of that start on page 163 in your book.
	Pupil:	Miss C. I'm putting that . . . (indistinguishable)
13.	1 C:	All right.
	Pupil:	At first look it.
14.	3 C:	What happened there?
	Pupil:	(indistinguishable)
15.	5 C:	What, what happened there? Tell about it. What happened?
	Pupil:	It increased a lot.
16.	(*a*) 3 C:	(*a*) Yes, she, she got much heavier that year didn't she? (Pupil: Yes.) She gained a lot of weight.
	(*b*) 7	(*b*) I don't know why.
	Pupil:	Because you made it that way.
17.	(*a*) 7 C:	(*a*) No, I took these right out of
	(*b*) 3	the records. You want to see it it's in the book. I didn't make it that way it was. (*b*) Which girl you got?
	Pupil:	B
		(Pupil comments are heard in far background.)
18.	3 C:	This one, no. It must be right here. See it goes up.

Pupil: Hmmm. She must have increased that much.

19. (*a*) 6 C: (*a*) Now you see, this shows it
(*b*) 5 better than your bar graph, (*b*) so while a bar graph's all right I would recommend that at least you fix the corners up as you suggested so that it looks like a line graph.

Pupil: (indistinguishable)

20. 3 C: Yes but this is, this is for the other one isn't it?

Pupil: Hummm.

21. (*a*) 7 C: (*a*) I'm saying that the bar graph
(*b*) 5 makes it hard to see really what happened. (b) Huh? Why don't you make a line graph?

TRANSCRIPT B

This transcript was prepared from stenographic notes of a teacher's verbalizations made during a typical class session. Statements made by pupils were not recorded by the stenographer. The explanatory comments noted within parentheses were taken from notes written by two observers during the same class session.

This sample of a teacher's verbal behavior was obtained during a short visit to a sixth-year classroom. It was the opinion of the observers that this teacher represented a "traditional" viewpoint in educational practice. There appeared to be considerable emphasis on recitation, deportment, and ready obedience to verbal instructions. The psychological intent of many of this teacher's statements may be unclear to the reader, but in the classroom situation intent was made apparent to the observers by means of accompanying gestures, inflections, and intonations. Many of the teacher's comments were sarcastic and "biting."

Approximately 25 pupils were present at the class session. Desks of the single and double variety were arranged in rows facing a blackboard which stretched across one end of the room. It was customary during this session for the teacher to walk back and forth in front of the rows of desks while conducting the class activities. The learning problem presented by the teacher during the observation concerned the development of the ability to tell time and to express time in writing. A large cardboard clock with adjustable hands hung from the blackboard and several similar clocks of the same construction were in the hands of the seated pupils. The teacher placed the hands of the clock to indicate "1:12" and had just asked for volunteers to write on the blackboard the time indicated by the clock.

Category

1. 5 "A!M!B! how about you?" (motioning toward the blackboard) (L. gingerly raises a hand)
2. 5 "L? All right, go ahead L., try it." (L., A., M., and B. all go to the blackboard)
3. 4 "Who else wants to try it?" (R. goes to the blackboard)
4. 6 (to L.) "We had 1:12, you wrote 1:00."
5. 5 "M! You try it . . . (M. begins to write on the board.) Let me see." (Teacher turns to R. who has been writing and erasing his work.)
6. 7 "R! *I* will give you just one more chance."
7. 5 "E! Go fix the clock again . . . the way you had it set before." (E. sets the clock to 11:30)
8. 3 "You see it's not twelve o'clock yet R."
9. 6 (annoyed) "Please put your hand down E."
10. 6 "All right, let's let A. have a chance and after that you pay a fine . . ." (A. writes on the board)
11. (*a*) 5 (*a*) "Leave that eleven . . . don't you
(*b*) 1 move it A. (*b*) It's wonderful."
12. (*a*) 3 (to A.) "(*a*) How many minutes in an
(*b*) 3 hour . . . (*b*) How many less 60 is it? . . . (A. responds correctly to both questions)
13. 1 (to A.) "Good girl! We congratulate you."
14. 1 (to M.) "You too! Good girl."
15. 4 "Now I am going to separate you." (Here the teacher was apparently referring to a frequently used device for dividing the class into two groups. The groups had titles paralleling those of well-known professional baseball teams.)
16. 5 "R! Would you get some chairs!"
17. 5 "D! You go get some chairs too!"
18. 4 "You may use these two tables . . ."
19. 5 "S! Pass the paper!"
20. 4 "Please don't take paper unless you're in the Giant group."
21. 5 (to D. who is carrying a chair) "Now you can bring it out." (The separation of the class was conducted with the speed and precision of military drill. One group remained in their seats while the other occupied chairs around tables at one side of the classroom. The basis for the division of the class was not apparent to the observers.)
22. 5 (to group at the tables working on arithmetic problems which were apparently being continued from the preceding day) "Now let me see, if I count three my Giants will be very busy . . . one . . ."
23. 5 (to group remaining in their seats)

"Please, Yankees put your books away . . ."

24. 5 "M! Move over there please."
25. 4 (absently to herself) "We need a few more chairs."
26. 5 (to Yankees) "All right now, these people move out this way . . ." (motioning) "One, two . . ."
27. 6 "E! in your own seat." (sharply)
28. 6 "Come on M., if you want to remain in this group, you have to be smart and not remain in your chair."
29. 4 "Now we are going to see how we do today."
30. 6 (to Giants, admonishing them) "All right, now you're going to be all by yourselves and you're going to be very smart."
31. 3 (to Yankees) "Now let me see if you can set the clock to the time we are supposed to be in school . . ."
32. 5 "M! You come and set the big clock."
33. 5 (helps M. adjust the hour hand) "Let's get this a little bit further down here . . ."
34. 6 (to M.) "You're five minutes late (for school)." (M. adjusts the clock)
35. 6 "Now you're too early!" (M. continues to manipulate the clock hands)
36. 1 "Now you're right." (R. takes over the hands of the clock)
37. 6 (to R.) "And you're late . . . now you're terribly late."
38. 6 (to R., sarcastically) "We will have to send you to the office."
39. 6 (to Giants who are whispering) "I'm afraid somebody will have to come out of that group."

REFERENCES

1. ANDERSON, H. H., & BREWER, J. E. *Studies of teachers' classroom personalities. App. Psychol. Monogr.*, 1946, No. 8. Stanford: Stanford Univer. Press.
2. BARTLETT, M. S. The use of transformations. *Biometrics*, 1947, **3**, 39-52.
3. COCHRAN, W. G. Some consequences when the assumptions for the analysis of variance are not satisfied. *Biometrics*, 1947, **3**, 22-38.
4. FISHER, R. A., & YATES, F. *Statistical tables for biological, agricultural and medical research.* (3rd Ed.) New York: Hafner, 1949.
5. FLANDERS, N. A. Personal-social anxiety as a factor in experimental learning situations. Unpublished doctoral dissertation, Univer. of Chicago, 1949.
6. KENDALL, M. G. *Rank correlation methods.* London: Griffin, 1948.
7. LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
8. MITZEL, H. E. Interest factors predictive of teachers' rapport with pupils. Unpublished doctoral dissertation, Univer. of Minnesota, 1952.
9. ORLEANS, J. S., *et al.* Some preliminary thoughts on the criteria of teacher effectiveness. *J. educ. Res.*, 1952, **45**, 641-648.
10. PERKINS, H. The effects of social-emotional climate and curriculum on group learning of in-service teachers. Unpublished doctoral dissertation, Univer. of Chicago, 1949.
11. RABINOWITZ, W., & TRAVERS, R. M. W. Problems of defining and assessing teacher effectiveness. *Educ. Theory*, 1953, **3**, 212-219.
12. THELEN, H. A. Experimental research toward a theory of instruction. *J. educ. Res.*, 1951, **45**, 89-136.
13. THELEN, H. A., & WITHALL, J. G. Three frames of reference: the description of climate. *Hum. Relat.*, 1949, **2**, 159-176.
14. WITHALL, J. G. The development of a technique for the measurement of social-emotional climate in classrooms. Unpublished doctoral dissertation, Univer. of Chicago, 1948.
15. WITHALL, J. G. The development of a technique for the measurement of social-emotional climate in classrooms. *J. exp. Educ.*, 1949, **17**, 347-361.
16. WITHALL, J. G. Assessment of the social-emotional climates experienced by a group of seventh graders as they moved from class to class. *Educ. psychol. Measmt*, 1952, **12**, 440-451.

(Accepted for publication July 27, 1953)